

What is Claimed is:

1. A method for identifying related ions in an input data set produced by analyzing a sample comprising:

5 correlating each row of data in an input data set with every other row of data in said input data set producing a correlation matrix, each row representing intensities over time for a particular mass to charge (m/z) range, each element of said correlation matrix including a correlation value and having associated row and column identifiers

identifying which rows in said input data set are associated with said correlation value;

10 clustering said correlation matrix identifying at least one group and at least one row of said correlation matrix as being in said at least one group, each group representing covarying chromatograms;

selecting at least one time period of interest for each group; and

producing a resultant spectrum for each group by sampling chromatograms

15 included in each of said groups at each of said at least one time period of interest of using a form of said input data set.

2. The method of Claim 1, further comprising:

filtering said input data set prior to performing said correlation.

20

3. The method of Claim 1, wherein said input data set includes only one of:
unfragmented spectrum, fragmented spectrum, and alternating unfragmented and
fragmented spectrum.

5 4. The method of Claim 3, wherein said input data set includes only alternating
unfragmented and fragmented spectra, and the method further comprising:

forming a combined spectrum including an unfragmented spectrum and related
fragmented spectrum;

performing said correlating, said clustering, said selecting and said producing
10 using said combined spectrum;

determining m/z values in said combined spectrum;

sampling said unfragmented spectra at said m/z values in said combined spectrum
at a scan maximum identified for said combined spectrum; and

determining that a sampled m/z value in said combined spectra is associated with
15 a parent if there is an intensity at said sampled m/z value, and determining that said
sampled m/z value in said combined spectra is associated with a fragment in an absence
of a signal at said sampled m/z value.

5. The method of Claim 3, wherein said input data set includes only alternating unfragmented and fragmented spectra, and the method further comprising:

performing said correlating, said clustering, said selecting and said producing using each of said unfragmented spectrum and said fragmented spectrum separately; and

5 matching each parent of said unfragmented spectrum to related fragments in said fragmented spectrum by determining which of said related fragments covary with said parent.

6. The method of Claim 1, wherein said clustering further comprises:

10 determining a first row of said correlation matrix including an element having a maximum correlation value of all correlation values in the correlation matrix being considered as candidates to be grouped;

determining a time scan associated with said first row;

for each element of said first row corresponding to a unique pairing of a row "i" and column "j", determining if a correlation value is greater than a predetermined value
15 and determining if a scan number at which row "j" maximizes is within a threshold number of scans of said time scan associated with said first row; and

if said each element is greater than said predetermined value and if a scan number at which row "j" maximizes is within a threshold number of scans of said time scan
20 associated with said first row, adding a row of said input data set to a current group wherein the row added has a row number equal to that of a column index "j" associated

with said each element, and excluding the row added from further consideration as one of said candidates for grouping.

7. The method of Claim 6, further comprising:

5 performing said determining a first row if a correlation value is greater than a predetermined value and if a scan number at which row “j” maximizes is within a threshold number of scans of said time scan associated with said first row for each element of the first row having an associated column index greater than an index associated with said first row.

10

8. The method of Claim 7, further comprising:

stopping formation of groups by said clustering when said maximum correlation value is less than a predetermined value.

15 9. The method of Claim 8, further comprising:

forming a new group with a selection of a subsequent row including an element having a maximum correlation value of all correlation values in the correlation matrix being considered as candidates to be grouped.

20

10. The method of Claim 3, wherein said input data set includes only alternating unfragmented and fragmented spectra, said input data set includes at least two components eluting at a same time and having a same elution profile, and the method
5 further comprising:

combining adjacent fragmented and unfragmented spectra resulting in a new combined data set including half the number of spectra in comparison to a total of spectra of said fragmented and unfragmented spectra;

producing a first resulting spectrum and a second resulting spectrum, said first
10 resulting spectrum corresponding to said unfragmented spectrum at a selected point in time and said second resulting spectrum corresponding to said fragmented spectrum at said selected point in time; and

performing processing to identify which of said at least two components is a parent associated with at least one fragment included in said fragmented spectrum.

15

11. The method of Claim 3, wherein said input data set includes only unfragmented spectrum, and said at least one group formed by said clustering identifies charge states and isotopes of a single component that coelute at a same time.

20 12. The method of Claim 3, wherein said input data set includes only fragmented spectrum, and said at least one group formed by said clustering identifies charge states, isotopes, and fragments of a single component that coelute at a same time.

13. The method of Claim 1, wherein said selecting time periods of interest includes:

summing intensities of extracted chromatograms for each group at each scan

5 point; and

determining a maximum intensity for each group at a particular scan point; and

wherein said producing a resultant spectrum includes:

sampling extracted chromatograms of each group at said particular scan point.

10 14. The method of Claim 1, wherein said input data set is produced using a mass spectrometer analyzing the sample.

15 15. The method of Claim 1, wherein said input data set includes at least one multimodal peak of an extracted ion chromatogram, a number of peaks in said

multimodal peak being represented as “n”, and the method further comprising:

determining at least one split point in said multimodal peak to divide said multimodal peak into portions;

apportioning a first row of said input data set corresponding to said multimodal peak into row portions in accordance with said at least one split point;

20 creating an additional “n-1” rows of data included in said input data set, each of said additional rows including a different one of said row portions;

removing from said first row all row portions included in said additional rows;

and

filling remaining elements of each of said additional rows and said first row.

5 16. The method of Claim 1, further comprising:

filtering said input data set producing a filtered data set, and wherein said form of
said input data set is said filtered data set.

10 17. The method of Claim 10, wherein at least two ions are two parent ions co-
eluting at a same time having a same elution profile and covary, and the method further
comprising:

performing other processing steps to associate each of said two parent ions with
corresponding fragment ions.

15 18. The method of Claim 10, wherein said at least two components are parent
peptides that coelute at a same time and exhibit similar elution profiles, and the method
further comprising:

determining that additional processing is needed to match each of said at least two
parent peptides with associated child fragments; and

20 performing said additional processing.

19. The method of Claim 10, wherein said at least two components are peptides.

20. A method for quantifying at least one ion in an input data set produced by analyzing a sample comprising:

correlating each row of data in an input data set with every other row of data in said input data set producing a correlation matrix, each row representing intensities over
5 time for a particular mass to charge (m/z) range, each element of said correlation matrix including a correlation value and having associated row and column identifiers identifying which rows in said input data set are associated with said correlation value;

clustering said correlation matrix identifying at least one group and at least one row of said correlation matrix as being in said at least one group, each group representing
10 chemically related components exhibiting correlated chromatographic behavior;

selecting at least one time period of interest for each group; and

producing a resultant spectrum for each group by sampling chromatograms included in each of said groups at each of said at least one time period of interest of using a form of said input data set.

15

21. A computer program product for identifying related ions in an input data set produced by analyzing a sample comprising:

machine executable code that correlates each row of data in an input data set with every other row of data in said input data set producing a correlation matrix, each row
5 representing intensities over time for a particular mass to charge (m/z) range, each element of said correlation matrix including a correlation value and having associated row and column identifiers identifying which rows in said input data set are associated with said correlation value;

machine executable code that clusters said correlation matrix identifying at least
10 one group and at least one row of said correlation matrix as being in said at least one group, each group representing covarying chromatograms;

machine executable code that selects at least one time period of interest for each group; and

machine executable code that produces a resultant spectrum for each group by
15 sampling chromatograms included in each of said groups at each of said at least one time period of interest of using a form of said input data set.

22. The computer program product of Claim 21, further comprising:

machine executable code that filters said input data set prior to performing said
20 correlation.

23. The computer program product of Claim 21, wherein said input data set includes only one of: unfragmented spectrum, fragmented spectrum, and alternating unfragmented and fragmented spectrum.

5 24. The computer program product of Claim 23, wherein said input data set includes only alternating unfragmented and fragmented spectra, and the computer program product further comprising:

 machine executable code that forms a combined spectrum including an unfragmented spectrum and related fragmented spectrum and wherein said machine
10 executable code that correlates, clusters, selects and produces uses said combined spectrum;

 machine executable code that determines m/z values in said combined spectrum;

 machine executable code that samples said unfragmented spectra at said m/z values in said combined spectrum at a scan maximum identified for said combined
15 spectrum; and

 machine executable code that determines that a sampled m/z value in said combined spectra is associated with a parent if there is an intensity at said sampled m/z value, and determines that said sampled m/z value in said combined spectra is associated with a fragment in an absence of a signal at said sampled m/z value.

20 25. The computer program product of Claim 23, wherein said input data set includes only alternating unfragmented and fragmented spectra, said machine executable

code that correlates, clusters, selects and produces uses each of said unfragmented spectrum and said fragmented spectrum separately; and the computer program product further comprising:

machine executable code that matches each parent of said unfragmented spectrum to related fragments in said fragmented spectrum by determining which of said related fragments covary with said parent.

26. The computer program product of Claim 21, wherein said clustering further comprises:

10 machine executable code that determines a first row of said correlation matrix including an element having a maximum correlation value of all correlation values in the correlation matrix being considered as candidates to be grouped;

machine executable code that determines a time scan associated with said first row;

15 machine executable code that, for each element of said first row corresponding to a unique pairing of a row "i" and column "j", determines if a correlation value is greater than a predetermined value and determining if a scan number at which row "j" maximizes is within a threshold number of scans of said time scan associated with said first row; and

machine executable code that, if said each element is greater than said
20 predetermined value and if a scan number at which row "j" maximizes is within a threshold number of scans of said time scan associated with said first row, adds a row of said input data set to a current group wherein the row added has a row number equal to

that of a column index “j” associated with said each element, and excluding the row added from further consideration as one of said candidates for grouping.

27. The computer program product of Claim 26, further comprising:

5 machine executable code that determines said first row if a correlation value is greater than a predetermined value and if a scan number at which row “j” maximizes is within a threshold number of scans of said time scan associated with said first row for each element of the first row having an associated column index greater than an index associated with said first row.

10

28. The computer program product of Claim 27, further comprising:

machine executable code that stops formation of groups by said clustering when said maximum correlation value is less than a predetermined value.

15

29. The computer program product of Claim 28, further comprising:

machine executable code that forms a new group with a selection of a subsequent row including an element having a maximum correlation value of all correlation values in the correlation matrix being considered as candidates to be grouped.

20

30. The computer program product of Claim 23, wherein said input data set includes only alternating unfragmented and fragmented spectra, said input data set includes at least two components eluting at a same time and having a same elution profile, and the computer program product further comprising:

5 , machine executable code that combines adjacent fragmented and unfragmented spectra resulting in a new combined data set including half the number of spectra in comparison to a total of spectra of said fragmented and unfragmented spectra;

 machine executable code that produces a first resulting spectrum and a second resulting spectrum, said first resulting spectrum corresponding to said unfragmented
10 spectrum at a selected point in time and said second resulting spectrum corresponding to said fragmented spectrum at said selected point in time; and

 machine executable code that performs processing to identify which of said at least two components is a parent associated with at least one fragment included in said fragmented spectrum.

15

31. The computer program product of Claim 23, wherein said input data set includes only unfragmented spectrum, and said at least one group formed by said clustering identifies charge states and isotopes of a single component that coelute at a same time.

20

32. The computer program product of Claim 23, wherein said input data set includes only fragmented spectrum, and said at least one group formed by said clustering

identifies charge states, isotopes, and fragments of a single component that coelute at a same time.

33. The computer program product of Claim 21, wherein said selecting time
5 periods of interest includes:

machine executable code that sums intensities of extracted chromatograms for each group at each scan point; and

machine executable code that determines a maximum intensity for each group at a particular scan point; and wherein said producing a resultant spectrum includes:

10 machine executable code that samples extracted chromatograms of each group at said particular scan point.

34. The computer program product of Claim 21, wherein said input data set is produced using a mass spectrometer analyzing the sample.

15

35. The computer program product of Claim 21, wherein said input data set includes at least one multimodal peak, wherein said input data set includes at least one multimodal peak of an extracted ion chromatogram, a number of peaks in said multimodal peak being represented as “n”, and the computer program product further
20 comprising:

machine executable code that determines at least one split point in said multimodal peak to divide said multimodal peak into portions;

machine executable code that apportions a first row of said input data set corresponding to said multimodal peak into row portions in accordance with said at least one split point;

machine executable code that creates an additional “n-1” rows of data included in said input data set, each of said additional rows including a different one of said row portions;

machine executable code that removes from said first row all row portions included in said additional rows; and

machine executable code that fills remaining elements of each of said additional rows and said first row.

36. The computer program product of Claim 21, further comprising:

machine executable code that filters said input data set producing a filtered data set, and wherein said form of said input data set is said filtered data set.

15

37. The machine executable code that of Claim 30, wherein at least two ions are two parent ions co-eluting at a same time having a same elution profile and covary, and the computer program product further comprising:

machine executable code that performs other processing steps to associate each of said two parent ions with corresponding fragment ions.

38. The computer program product of Claim 30, wherein said at least two components are parent peptides that coelute at a same time and exhibit similar elution profiles, and the computer program product further comprising:

machine executable code that determines that additional processing is needed to

5 match each of said at least two parent peptides with associated child fragments; and

machine executable code that performs said additional processing.

39. The computer program product of Claim 29, wherein said at least two components are peptides.

10

40. A computer program product for quantifying at least one ion in an input data set produced by analyzing a sample comprising:

machine executable code that correlates each row of data in an input data set with every other row of data in said input data set producing a correlation matrix, each row
5 representing intensities over time for a particular mass to charge (m/z) range, each element of said correlation matrix including a correlation value and having associated row and column identifiers identifying which rows in said input data set are associated with said correlation value;

machine executable code that clusters said correlation matrix identifying at least
10 one group and at least one row of said correlation matrix as being in said at least one group, each group representing chemically related components exhibiting correlated chromatographic behavior;

machine executable code that selects at least one time period of interest for each group; and

15 machine executable code that produces a resultant spectrum for each group by sampling chromatograms included in each of said groups at each of said at least one time period of interest of using a form of said input data set.